Asymptotic tests for general linear hypotheses on variance components in models of commutative quadratic type

Joachim Hartung and Guido Knapp
Department of Statistics, University of Dortmund, D–44221 Dortmund, Germany

Abstract: In this paper we derive asymptotic χ^2–tests for general linear hypotheses on variance components using repeated variance components models. In two examples, the two–way nested classification model and the two–way crossed classification model with interaction, we explicitly investigate the properties of the asymptotic tests in small sample sizes.

Key Words: Wald– and likelihood ratio test statistic, repeated variance components model, linear hypotheses on variance components

1 Introduction

In this paper we consider linear hypotheses on variance components as

$$H_0: \; K\sigma = d \quad , \tag{1}$$

where $\sigma = (\sigma_1^2, \ldots, \sigma_m^2)^T$ denotes a vector of unknown variance components, K is a known $(p \times m)$–matrix with rk $(K) = p \leq m$ and $d \in I\!\!R^p$ a known constant. For special linear combinations of variance components exact F– and χ^2–tests can be derived and in El–Bassiouni and Seely (1980) it is shown that under certain circumstances these tests are uniformly most powerful unbiased. However, no exact tests for example are known for testing that the variance of a certain factor is equal to a given $d_1 > 0$ or that the difference between two variance components equals a certain value. Here, we develop asymptotic χ^2–tests for such hypotheses.

In section 2 we consider the class of variance components models of commutative quadratic type (see e. g. Seely (1971), Elpelt (1989), Hartung (1981, sec. 5)) and introduce repeated variance components models (cf. Brown (1976)). Then, Wald and likelihood ratio test statistics are derived in section 3 using the approach of repeated variance components model, where the asymptotic results refer to a 'large' number of observations in the experimental designs which can be interpreted as several independent observations from a reduced design. In practice, however, we deal with non–repeated models. Thus, in section 4 we consider, as examples, the two–way nested classification model and the two–way crossed classification model with interaction, where we explicitly study hypotheses about the differences of two variance components. In both models wn can directly give the Wald test statistics for the hypotheses as well as the unrestricted maximum likelihood estimators. In the two–way nested classification model, however, we use a numerical algorithm to maximize the likelihood function under the hypothesis, whereas in the two–way crossed classification model with interaction we give in addition an explicit approximation of the likelihood ratio test statistic which does not need a numerical algorithm. In a simulation study we examine the finite properties of the derived tests, especially in situations where the sample sizes are really 'small' and show that the asymptotic works satisfactorily in these cases. Hereby, a clear preference of the likelihood ratio test can be stated, on the whole.

Throughout this paper we use the following notation. For a real matrix A let A^T denote the transposed, A^+ the Moore–Penrose–inverse, rk (A) the rank, tr (A) the trace, and $\mathcal{R}(A)$ the range of A. Further we denote by $\otimes$ the Kronecker product, by I_r the $(r \times r)$–identity matrix, by 1_r the vector of r ones, by $J_r = 1_r 1_r^T$ the $(r \times r)$–matrix of ones, and by $0_{s \times t}$ the $(s \times t)$–matrix of zeros.

2 The repeated variance components model

We consider a q–dimensional observable random vector, say Y, that follows the general linear variance components model

$$\mathrm{E}\,(Y) = X\beta \qquad \text{and} \qquad \mathrm{Cov}\,(Y) = \sum_{i=1}^{m} \sigma_i^2\, U_i \;, \tag{2}$$

where the $(q \times l)$–matrix X and the symmetric positive semidefinite $(q \times q)$–matrices $U_1, \ldots, U_m$ are known, whereas the parameter vector β varies in $\mathbb{R}^l$ and the parameter vector $\sigma = (\sigma_1^2, \ldots, \sigma_m^2)^T$ varies in Ω, a subset of $\mathbb{R}^m_{(+)}$, the nonnegative orthant of $\mathbb{R}^m$. The variance component σ_m^2 is assumed to be strictly positive and U_m is positive definite to ensure the positive definiteness of the variance–covariance matrix..

In this model we make the following assumptions.

Assumption 1: The random vector Y has a q–dimensional normal distribution.
Assumption 2: $\Psi = \mathit{span}\{XX^+, U_1, \ldots, U_m\}$ forms a $(m+1)$–dimensional commutative quadratic subspace of all real symmetric $(q \times q)$–matrices, i. e. Ψ is a subspace and $A, B \in \Psi$ implies $A^2 \in \Psi$ and $AB = BA$.

By lemma 6 in Seely (1971) there exists a basis $P_0, P_1, \ldots, P_m$ of Ψ with $P_0 = XX^+$, where P_i, $i = 0, 1, \ldots, m$, is idempotent and $P_i P_j = 0$, $i \neq j$. Then there is a nonsingular $((m+1) \times (m+1))$–matrix

$$\Lambda^* = ((\lambda_{ij})_{i,j=0,1,\ldots,m}) \tag{3}$$

such that

$$U_i = \sum_{j=0}^{m} \lambda_{ij} P_j \;,\; i = 0, 1, \ldots, m, \quad \text{with} \quad U_0 = P_0 = XX^+ \;. \tag{4}$$

Assumption 3: For all $j = 1, \ldots, m$ and $\sigma \in \Omega$ it holds

$$\tau_j = \sum_{j=1}^{m} \lambda_{ij}\, \sigma_i^2 > 0 \;. \tag{5}$$

The vector $\tau = (\tau_1, \dots, \tau_m)^T$ can be expressed as

$$\tau = \Lambda^T \sigma \ , \tag{6}$$

where the nonsingular $(m \times m)$–matrix Λ is the submatrix of Λ^*, which results from deleting the first row and the first column of Λ^*.

The three assumptions are fulfilled, for example, in the balanced variance components models if, as it is usually assumed, the residual variance is strictly positive.

Let us consider the quadratic forms

$$T_j = Y^T P_j Y / f_j \ , \ j = 1, \dots, m, \tag{7}$$

where tr $(P_j) = f_j$, $j = 1, \dots, m$. In model (**??**) it holds that these quadratic forms are stochastically independent and that $f_j \cdot T_j/\tau_j$, $j = 1, \dots, m$, is central χ^2–distributed with f_j degrees of freedom. It follows that the expectation vector of the random vector $T = (T_1, \dots, T_m)^T$ is given by the vector τ and the variance–covariance matrix of T is

$$D(\sigma) = 2 \cdot \mathrm{diag}\,(\tau_1^2, \dots, \tau_m^2) \ . \tag{8}$$

Now the model (**??**) is ν–times statistically independently repeated, i. e. we observe independent random vectors Y_ρ, $\rho = 1, \dots, \nu$, which have the same distributional properties as Y from (**??**). Thus, we get the following model

$$\begin{gathered} \tilde{Y} = (Y_1^T, Y_2^T, \dots, Y_\nu^T)^T \\ \text{with} \quad \mathrm{E}\,(\tilde{Y}) = (1_\nu \otimes X)\beta \ , \ \mathrm{Cov}\,(\tilde{Y}) = \sum_{i=1}^m \sigma_i^2 \, (I_\nu \otimes U_i) \ . \end{gathered} \tag{9}$$

Due to (**??**) the variance–covariance matrix of $\tilde{Y}$ can be expressed as

$$\mathrm{Cov}\,(\tilde{Y}) = \sum_{i=1}^m \sigma_i^2 \, (I_\nu \otimes U_i) = \sum_{j=0}^m \tau_j \, (I_\nu \otimes P_j) \ , \tag{10}$$

where τ_0 is a linear combination of $\tau_1, \dots, \tau_m$.

Note that in the repeated model (**??**) a corresponding property like the assumption 2 in the non–repeated model (**??**) does not hold, so that we cannot refer to the results from Seely (1971).

Lemma 1:

In model (**??**) it holds

i) the $m+1$ matrices $(I_\nu \otimes P_i)$, $i = 0, 1, \ldots, m$, are idempotent and mutually orthogonal matrices;

ii) the $m+1$ statistics

$$(I_\nu \otimes P_0)\tilde{Y} \ , \ \tilde{Y}^T (I_\nu \otimes P_j)\tilde{Y} \ , \ j = 1, \ldots, m \ , \tag{11}$$

are minimal sufficient statistics for this model.

Proof:

i) It is $(I_\nu \otimes P_i)^2 = I_\nu \otimes P_i^2 = I_\nu \otimes P_i$, $i = 0, 1, \ldots, m$

and $(I_\nu \otimes P_i)(I_\nu \otimes P_j) = I_\nu \otimes P_i P_j = I_n \otimes 0_{q \times q}$, $i \neq j$,

because the $m+1$ matrices P_i, $i = 0, 1, \ldots, m$, are idempotent and mutually orthogonal.

ii) Due to i) the inverse of the variance–covariance matrix Cov $(\tilde{Y})$ can be written as

$$\text{Cov } (\tilde{Y})^{-1} = \sum_{j=0}^{m} \frac{1}{\tau_j} (I_\nu \otimes P_j). \tag{12}$$

Using the fact that $P_j X = 0_{q \times l}$, $j = 1, \ldots, m$, it holds for the argument of the exponential function in the density of the $\nu \cdot q$–dimensional normal distribution

$$\begin{aligned} & (\tilde{Y} - (1_\nu \otimes X)\beta)^T \ (\text{Cov } (\tilde{Y}))^{-1} \ (\tilde{Y} - (1_\nu \otimes X)\beta) \qquad (13) \\ = \ & \frac{1}{\tau_0} (\tilde{Y} - (1_\nu \otimes X)\beta)^T (I_\nu \otimes P_0)(\tilde{Y} - (1_\nu \otimes X)\beta) + \sum_{j=1}^{m} \frac{1}{\tau_j} \tilde{Y}^T (I_\nu \otimes P_j) \tilde{Y} \ . \end{aligned}$$

Using the factorization criterion (see e. g. Graybill (1976), Theorem 2.6.1, p. 69) it follows from (**??**) that the statistics $(I_\nu \otimes P_0)\tilde{Y}$, $\tilde{Y}^T (I_\nu \otimes P_j)\tilde{Y}$, $j = 1, \ldots, m$, are sufficient statistics. With Theorem 2.6.2 in Graybill (1976, p. 70) it can be stated that the set of the above statistics is a set of minimal sufficient statistics.

Lemma 2:

i) The GLS estimator of the parameter vector β in model (**??**) depends only on the sufficient statistic $(I_\nu \otimes P_0)\tilde{Y}$.

ii) Quadratic unbiased estimators of $\tau_1, \ldots, \tau_m$ are based on the sufficient statistics $\tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}$, $j = 1, \ldots, m$.

Proof:

i) With $P_j X = 0_{q\times l}$, $j = 1, \ldots, m$ and (**??**) the term $(1_\nu \otimes X)^T \,(\mathrm{Cov}\,(\tilde{Y}))^{-1}\, \tilde{Y}$ reduces to $\frac{1}{\tau_0}(1_\nu \otimes X)^T(I_\nu \otimes P_0)\tilde{Y}$, so that the GLS estimator of β has the form
$\hat{\beta} = \big((1_\nu \otimes X)^T(I_\nu \otimes P_0)(1_\nu \otimes X)\big)^+(1_\nu \otimes X)^T(I_\nu \otimes P_0)\tilde{Y}$.

ii) The expected value of $\tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}$, $j = 1, \ldots, m$, is given by
$\mathrm{E}\,(\tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}) = \mathrm{tr}\,(I_\nu \otimes P_j)\mathrm{Cov}\,(\tilde{Y}) = \tau_j \cdot \mathrm{tr}\,(I_\nu \otimes P_j)$,
because $P_j X = 0_{q\times l}$, $j = 1, \ldots, m$. Division by the trace of the projection matrix yields the result. This completes the proof.

Due to (**??**) we have a unique relation between τ and σ. Hence, we use the quadratic forms $\tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}$, $j = 1, \ldots, m$, to make inference about the unknown vector of variance components.

3 Derivation of the test statistics

In model (**??**) we consider the quadratic forms $\tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}$, $j = 1, \ldots, m$, and define for all $j = 1, \ldots, m$

$$T_j^\nu = \tilde{Y}^T(I_\nu \otimes P_j)\tilde{Y}/(\nu \cdot f_j) = \frac{1}{\nu}\sum_{\rho=1}^{\nu} Y_\rho^T P_j Y_\rho / f_j = \frac{1}{\nu}\sum_{\rho=1}^{\nu} T_j^\rho \ , \tag{14}$$

where $T_j^\rho = Y_\rho^T P_j Y_\rho / f_j$, $\rho = 1, \ldots, \nu$, $j = 1, \ldots, m$.

Let us denote $T^\nu = (T_1^\nu, \ldots, T_m^\nu)^T$ then it holds

$$\mathrm{E}\,(T^\nu) = \Lambda^T\sigma = \tau \quad , \tag{15}$$

and the variance–covariance matrix of T^ν is a diagonal $(m \times m)$–matrix given by

$$D^\nu(\sigma) = 2 \cdot \mathrm{diag}\,(\tau_1^2/(\nu \cdot f_1), \ldots, \tau_m^2/(\nu \cdot f_m)) = D(\sigma)/v\ , \qquad (16)$$

and $D(\sigma)$ is the variance–covariance matrix of T from (**??**) in the corresponding non–repeated model (**??**).

For each $\rho = 1, \ldots, \nu$ the random variables $f_j \cdot T_j^\rho/\tau_j$, $j = 1, \ldots, m$, are independent χ^2–distributed random variables with f_j degrees of freedom. Thus, we consider the likelihood function

$$L(\sigma) = \prod_{\rho=1}^{\nu} \prod_{i=1}^{m} (C_i)^{-1} \left(\frac{f_i}{\tau_i}\right)^{f_i/2} (T_i^\rho)^{(f_i-2)/2} \exp\left\{ -\frac{1}{2} \frac{f_i \cdot T_i^\rho}{\tau_i} \right\} . \qquad (17)$$

with $(C_i)^{-1} = 2^{f_i/2}\, \Gamma(f_i/2)$, $i = 1, \ldots m$, and $\Gamma(x)$ denotes the gamma function.
So, the log–likelihood function reads

$$l(\sigma) = \sum_{\rho=1}^{\nu} \sum_{i=1}^{m} \left\{ \ln(C_i)^{-1} + \frac{f_i}{2} \ln\left(\frac{f_i}{\tau_i}\right) + \left(\frac{f_i - 2}{2}\right) \ln T_i^\rho - \frac{1}{2} \frac{f_i \cdot T_i^\rho}{\tau_i} \right\} . \qquad (18)$$

For the first derivatives of the log–likelihood function (**??**) we get

$$\begin{aligned} \frac{\partial l(\sigma)}{\partial \sigma_j^2} &= \sum_{\rho=1}^{\nu} \sum_{i=1}^{m} \left(\frac{f_i}{2\tau_i^2} \cdot \lambda_{ij} \cdot T_i^\rho - \frac{f_i}{2\tau_i} \cdot \lambda_{ij} \right) \\ &= \sum_{i=1}^{m} \left(\frac{\nu \cdot f_i}{2\tau_i^2} \cdot \lambda_{ij} \cdot T_i^\nu - \frac{\nu \cdot f_i}{2\tau_i} \cdot \lambda_{ij} \right)\ ,\ j = 1, \ldots, m, \end{aligned} \qquad (19)$$

so that

$$\frac{\partial l(\sigma)}{\partial \sigma} = \Lambda (D^\nu(\sigma))^{-1} (T^\nu - \Lambda^T \sigma) \quad . \qquad (20)$$

Due to (**??**) the maximum likelihood estimator of σ has the form

$$\hat{\sigma} = (\Lambda^T)^{-1} \cdot T^\nu \quad , \qquad (21)$$

and thus, the maximum likelihood estimator of τ is given by

$$\hat{\tau} = T^\nu \quad . \qquad (22)$$

The maximum likelihood estimator in (**??**) coincides with the usual ANOVA–estimator and asymptotically yields nonnegative estimates of the variance components (cf. Brown (1976)).

For the second derivatives of the log–likelihood function (**??**) we obtain

$$\begin{aligned}\frac{\partial^2 l(\sigma)}{\partial\sigma_j^2\partial\sigma_k^2} &= \sum_{\rho=1}^{\nu}\sum_{i=1}^{m}\Big(-\frac{f_i}{\tau_i^3}\lambda_{ij}\,\lambda_{ik}\,T_i^{\rho}+\frac{f_i}{2\tau_i^2}\lambda_{ij}\,\lambda_{ik}\Big)\\ &= \sum_{i=1}^{m}\Big(-\frac{\nu\cdot f_i}{\tau_i^3}\lambda_{ij}\,\lambda_{ik}\,T_i^{\nu}+\frac{\nu\cdot f_i}{2\tau_i^2}\lambda_{ij}\,\lambda_{ik}\Big)\ ,\ j,k=1,\dots,m, \qquad (23)\end{aligned}$$

so that the mean values of these derivatives are

$$\mathrm{E}\left(\frac{\partial^2 l(\sigma)}{\partial\sigma_j^2\partial\sigma_k^2}\right) = -\sum_{i=1}^{m}\frac{\nu\cdot f_i}{2\tau_i^2}\lambda_{ij}\,\lambda_{ik} \quad ,\ j,k=1,\dots,m\ . \qquad (24)$$

Thus, the information matrix is given by

$$\begin{aligned}I^{\nu}(\sigma)=\mathrm{E}\left(-\frac{\partial^2 l(\sigma)}{\partial\sigma\partial\sigma^T}\right) &= \Lambda(D^{\nu}(\sigma))^{-1}\Lambda^T \qquad (25)\\ &= \nu\cdot\Lambda(D(\sigma))^{-1}\Lambda^T=\nu\cdot I(\sigma)\quad ,\end{aligned}$$

where $I(\sigma)$ is the information matrix in the corresponding non–repeated model (**??**).

Due to the results of Anderson (1973) and Brown (1976), respectively, we can state the following theorem.

Theorem 1:
In model (**??**) it holds that $\sqrt{\nu}\,(\hat\sigma-\sigma)$ is asymptotically normally distributed with mean vector 0 and variance–covariance matrix $(\Lambda(D(\sigma))^{-1}\Lambda^T)^{-1}$ for $\nu\to\infty$.

Corollary 1:
Under the hypothesis $H_0: K\sigma=d$, $\sqrt{\nu}\,(K\hat\sigma-d)$ is asymptotically normally distributed with mean vector 0 and variance–covariance matrix $K(\Lambda(D(\sigma))^{-1}\Lambda^T)^{-1}K^T$ for $\nu\to\infty$.

Thus, the Wald–type test statistic for testing the general linear hypothesis (**??**) is given by

$$W=(K\hat\sigma-d)^T(K(\Lambda(D^{\nu}(\sigma))^{-1}\Lambda^T)^{-1}K^T)^{-1}(K\hat\sigma-d)\ , \qquad (26)$$

which is under H_0 asymptotically χ^2–distributed with rk (K) degrees of freedom (cf. Rao (1973), p. 188)). For an application of the Wald test a consistent estimator of σ, usually the maximum likelihood estimator $\hat{\sigma}$, has to be replaced in $D^\nu(\sigma)$.

An asymptotically equivalent test to the Wald test is given by the likelihood ratio test. Thus, we consider the ratio

$$\max_{\sigma\,:\;K\sigma=d} L(\sigma) \Big/ \max_{\sigma} L(\sigma). \tag{27}$$

Considering the Lagrangian function

$$\mathcal{L}(\sigma,\lambda) = l(\sigma) - \lambda^T(K\sigma - d), \tag{28}$$

the maximum likelihood estimator of σ under H_0, say $\bar{\sigma} = (\bar{\sigma}_1^2, \ldots, \bar{\sigma}_m^2)^T$, is a solution of

$$\begin{aligned} (T^\nu - \tau) - D^\nu(\sigma)\Lambda^{-1}K^T\lambda &= 0 \\ K\sigma &= d\,, \end{aligned} \tag{29}$$

where $\lambda \in I\!R^p$ is a vector of Lagrange multipliers.

Theorem 2:
The test statistic

$$LR = 2\big(l(\hat{\sigma}) - l(\bar{\sigma})\big) \tag{30}$$

is under $H_0 : K\sigma = d$ asymptotically χ^2–distributed with rk (K) degrees of freedom.

Proof: We note that the likelihood function (**??**) is built of independent identical distributed random vectors $T^\rho = (T_1^\rho, \ldots, T_m^\rho)^T$, $\rho = 1, \ldots \nu$, so that the proof is given using standard arguments of maximum likelihood theory (see e. g. Rao (1973), p. 418–419).

Using the representation of the log–likelihood function from (**??**) the likelihood ratio test statistic (**??**) can also be expressed as

$$LR = \sum_{i=1}^m \nu \cdot f_i \left\{ \frac{T_i^\nu}{\sum_{j=1}^m \lambda_{ij}\bar{\sigma}_i^2} - \ln\left(\frac{T_i^\nu}{\sum_{j=1}^m \lambda_{ij}\bar{\sigma}_i^2}\right) - 1 \right\}\;. \tag{31}$$

4 Two Examples

4.1 Two–way nested classification model

We consider the balanced two–way nested classification model with random effects given by

$$y_{ijk} = \mu + a_i + b_{ij} + e_{ijk} \qquad (32)$$
$$i = 1, \ldots, r;\ j = 1, \ldots, s;\ k = 1, \ldots, t;\ n = rst,$$

where $\mu \in I\!R$ is a fixed effect and $a_1, \ldots, a_r,\ b_{11}, \ldots, b_{rs},\ e_{111}, \ldots, e_{rst}$ are independent normally distributed random effects with $\mathrm{E}\,(a_i) = \mathrm{E}\,(b_{ij}) = \mathrm{E}\,(e_{ijk}) = 0$ and $\mathrm{Var}\,(a_i) = \sigma_a^2$, $\mathrm{Var}\,(b_{ij}) = \sigma_b^2$, $\mathrm{Var}\,(e_{ijk}) = \sigma_e^2 > 0$ for all i, j, and k , so $\sigma = (\sigma_a^2, \sigma_b^2, \sigma_e^2)^T$.

The unique basis of projection matrices in this model is given by

$$\begin{array}{ll} P_0 = \frac{1}{n} J_n & , \\ P_a = (I_r - \frac{1}{r} J_r) \otimes \frac{1}{st} J_{st} & ,\ \ \mathrm{tr}\,(P_a) = r - 1\ , \\ P_b = I_r \otimes (I_s - \frac{1}{s} J_s) \otimes \frac{1}{t} J_t & ,\ \ \mathrm{tr}\,(P_b) = r(s-1)\ , \\ P_e = I_{rs} \otimes (I_t - \frac{1}{t} J_t) & ,\ \ \mathrm{tr}\,(P_e) = rs(t-1)\ , \end{array} \qquad (33)$$

and the matrix Λ has the form

$$\Lambda = \begin{pmatrix} st & 0 & 0 \\ t & t & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad . \qquad (34)$$

With $y = (y_{111}, y_{112}, \ldots, y_{rst})^T$ let us denote the mean sum of squares of the random effects as

$$\begin{aligned} M_1 &= y^T P_a y/(r-1)\ , \\ M_2 &= y^T P_b y/r(s-1)\ , \\ M_3 &= y^T P_e y/rs(t-1)\ . \end{aligned} \qquad (35)$$

For an application of the Wald test statistic we have to replace $D(\sigma)$ in (**??**) by a consistent estimator. Using a result from Hartung and Voet (1986) the best invariant unbiased estimator for $D(\sigma)$ is given by

$$\widehat{D(\sigma)} = 2 \cdot \mathrm{diag}\left(\frac{M_1^2}{r+1}, \frac{M_2^2}{r(s-1)+2}, \frac{M_3^2}{rs(t-1)+2} \right) . \qquad (36)$$

For testing the hypothesis $H_0 : \sigma_a^2 = \sigma_b^2$ with $K = (1, -1, 0)^T$ and $d = 0$ the Wald test statistic has the form

$$W_1 = \frac{\left(\frac{1}{st}(M_1 - M_2) - \frac{1}{t}(M_2 - M_3)\right)^2}{\frac{1}{s^2 t^2}\left(\frac{2\,M_1^2}{r+1} + \frac{2\,(s+1)^2\,M_2^2}{r(s-1)+2} + \frac{2\,s^2\,M_3^2}{rs(t-1)+2}\right)}, \tag{37}$$

which can also be expressed in terms of the maximum likelihood estimators $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$ as

$$W_1 = \frac{(\hat{\sigma}_a^2 - \hat{\sigma}_b^2)^2}{\widehat{\mathrm{Var}}\,(\hat{\sigma}_a^2 - \hat{\sigma}_b^2)}. \tag{38}$$

The test statistic W_1 is under H_0 asymptotically χ^2–distributed with one degree of freedom.

In order to apply the likelihood ratio test statistic for testing $H_0 : \sigma_a^2 = \sigma_b^2$ we have to make use of a numerical algorithm to maximize the log-likelihood under H_0. In the following simulation studies, which have been carried out in SAS 6.12 using PROC IML, we use the Newton-Raphson ridge optimization algorithm to obtain the maximum likelihood estimator under H_0.

In the first simulation study we investigate the behaviour of the significance level and the power of both tests where we focus our attention on 'small' degrees of freedom of the mean sum of squares. Due to the fact that a given two–way nested classification model can possibly be interpreted as a replication of a reduced design, which depends on the the number r of levels of the A–factor, we only consider sample sizes with increasing r and three pairs of sample sizes (s, t) to make the simulations not too complex. For the error variance σ_e^2 we always choose the value one. In table 1 the result for the Wald and likelikood ratio test concerning the estimated size of the tests given the nominal level of $\alpha = 0.01$ and $\alpha = 0.05$, respectively, are presented based on 10 000 simulations of the model.

We observe that the estimated significance levels of the likelihood ratio test are nearly independent from the chosen sample sizes and mostly exceed the nominal significance levels but in a compatible manner; merely in five cases with $\alpha = 0.01$ the estimated sizes are somewhat smaller than 0.01. The estimated significance levels of the Wald test, however, do not show such a homogeneous behaviour as the estimated sizes of the likelihood ratio test. For all r with $s = t = 3$ the estimated sizes of the Wald test always considerably fall below the nominal significance level. Only the case $r = 20$ indicates that for larger r the actual size of the test reaches the nominal size. If we choose $s = 5$ and $t = 6$ the estimated sizes always exceed the nominal level except for $r = 3, 4, 5$ with $\alpha = 0.01$, but the differences are not so severe. Regarding the sample size $s = 8$ and $t = 10$ for all r the estimated significance levels are considerably larger than the nominal significance levels. Again the case $r = 20$ indicates that for larger r the actual size of the test may go towards the nominal significance level. Consequently, the likelihood ratio test seems to be preferable to the Wald test in small sample sizes.

Yet, we perform another simulation study to compare the power of both tests and we restrict to all sample sizes r with $s = 5$ and $t = 6$, because in these cases the estimated sizes of both tests are rather similar. As possible alternative hypotheses we consider the cases $\delta = \sigma_b^2 - \sigma_a^2 = 0.1, 0.25, 0.5, 0.75, 1$, and 2. The results of this simulation study are given in table 2, where again every estimated point of the power function is based on 10000 simulations. For all sample sizes the estimated significance level of the Wald test is always a little bit larger than the estimated significance level of the likelihood ratio test, so one may expect that the power function of the Wald test is also larger than the one of the likelihood ratio test. The simulation study corroborates this presumption, but the gain of the power of the Wald test is even much more intense than the one of the likelihood ratio test, i. e. in situations with comparable sizes of both tests the Wald test detects better deviations from the null hypothesis than the likelihood ratio test.

4.2 Two–way classification model with interaction

Let us consider the balanced two-way classification random model with interaction given by

$$\begin{gathered} y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}\ , \\ i = 1, \dots, r;\ j = 1, \dots, r;\ k = 1, \dots, t;\ n = rst, \end{gathered} \tag{39}$$

where $\mu \in I\!R$ is a fixed effect and $a_1, \dots, a_r$, $b_1, \dots, b_s$, $(ab)_{11}, \dots, (ab)_{rs}$, $e_{111}, \dots, e_{rst}$ are independent normally distributed random effects with $\mathrm{E}\,(a_i) = \mathrm{E}\,(b_j) = \mathrm{E}\,((ab)_{ij}) = \mathrm{E}\,(e_{ijk}) = 0$ and $\mathrm{Var}\,(a_i) = \sigma_a^2$, $\mathrm{Var}\,(b_j) = \sigma_b^2$, $\mathrm{Var}\,((ab)_{ij}) = \sigma_{ab}^2$, $\mathrm{Var}\,(e_{ijk}) = \sigma_e^2 > 0$ for all i, j and k, so $\sigma = (\sigma_a^2, \sigma_b^2, \sigma_{ab}^2, \sigma_e^2)^T$.

Let M_1, M_2, M_3 and M_4 represent the A–factor, B–factor, AB–interaction and residual error mean squares, it holds

$$\mathrm{E}(M) = \Lambda^T \sigma, \quad M = (M_1, M_2, M_3, M_4)^T, \tag{40}$$

and

$$\Lambda = \begin{pmatrix} st & 0 & 0 & 0 \\ 0 & rt & 0 & 0 \\ t & t & t & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{41}$$

where the unique basis of projection matrices is given by

$$\begin{array}{lcl} P_0 = \frac{1}{n} J_n & , & \\ P_a = \left(I_r - \frac{1}{r} J_r\right) \otimes \frac{1}{st} J_{st} & , & \mathrm{tr}\,(P_a) = r - 1\ , \\ P_b = \frac{1}{r} J_r \otimes \left(I_s - \frac{1}{s} J_s\right) \otimes \frac{1}{t} J_t & , & \mathrm{tr}\,(P_b) = s - 1\ , \\ P_{ab} = \left(I_r - \frac{1}{r} J_r\right) \otimes \left(I_s - \frac{1}{s} J_s\right) \otimes \frac{1}{t} J_t & , & \mathrm{tr}\,(P_{ab}) = (r-1)(s-1)\ , \\ P_e = I_{rs} \otimes (I_t - \frac{1}{t} J_t) & , & \mathrm{tr}\,(P_e) = rs(t-1)\ . \end{array} \tag{42}$$

The best invariant unbiased estimator of the covariance matrix $D(\sigma)$ (cf. Hartung, Voet 1986) has the form

$$\widehat{D(\sigma)} = 2 \cdot \mathrm{diag}\left(\frac{M_1^2}{r+1}, \frac{M_2^2}{s+1}, \frac{M_3^2}{(r-1)(s-1)+2}, \frac{M_4^2}{rs(t-1)+2}\right). \tag{43}$$

Thus, the Wald test statistic for testing the hypothesis $H_0 : \sigma_a^2 = \sigma_b^2$, with $K = (1,-1,0,0)^T$ and $d = 0$, can be described as

$$W_1 = \frac{\Big(r(M_1 - M_3) - s(M_2 - M_3)\Big)^2}{\dfrac{2r^2M_1^2}{r+1} + \dfrac{2s^2M_2^2}{s+1} + \dfrac{2(r-s)^2M_3^3}{(r-1)(s-1)+2}} \quad , \tag{44}$$

which is under H_0 asymptotically χ^2–distributed with one degree of freedom.

In this model we explicitly consider the equations (**??**) which has to be solved by the maximum likelihood estimator under H_0. Here, (**??**) has the form

$$\begin{array}{lcl} (M_1 - \tau_1) + 2\lambda \dfrac{\tau_1^2}{st(r-1)} & = & 0 \\ (M_2 - \tau_2) - 2\lambda \dfrac{\tau_2^2}{rt(s-1)} & = & 0 \\ (M_3 - \tau_3) + 2\lambda \left(\dfrac{1}{rt} - \dfrac{1}{st}\right) \dfrac{\tau_3^2}{(r-1)(s-1)} & = & 0 \\ (M_4 - \tau_4) & = & 0 \\ \frac{1}{s}\tau_1 - \frac{1}{r}\tau_2 + \left(\frac{1}{r} - \frac{1}{s}\right)\tau_3 & = & 0 \end{array} \tag{45}$$

If both main effects have the same number of levels, i. e. $r = s$, we get the following solution of (**??**)

$$\bar{\tau}_4 = M_4 \ , \ \bar{\tau}_3 = M_3 \ , \ \text{and } \bar{\tau}_2 = \bar{\tau}_1 = (M_1 + M_2)/2 \quad . \tag{46}$$

So the test statistic (**??**) can be written as

$$LR_1 = (r-1)\left\{\ln\left(\frac{M_1+M_2}{2M_1}\right) + \ln\left(\frac{M_1+M_2}{2M_2}\right)\right\} \ . \tag{47}$$

Under H_0, the mean value of LR_1 is given by

$$\mathrm{E}\,(LR_1) = 2(r-1)\Big\{\mathrm{E}\ \ln\left(\chi^2_{2(r-1)}\right) - \mathrm{E}\ \ln\left(\chi^2_{r-1}\right) - \ln 2\Big\} \, . \tag{48}$$

In Bartlett and Kendall (1946) it is shown that the mean value of the logarithm of a χ^2–distributed random variable with f degrees of freedom is given by

$$\mathrm{E}\ \ln(\chi^2_f) = \ln 2 + \psi(f/2), \tag{49}$$

where $\psi(x) = \mathrm{d}\ \ln\Gamma(x)/\mathrm{d}x$ is the psi function. With an approximation of the psi function given in Abramowitz and Stegun (1964, p. 259) we get the following approximation of the mean value

$$\mathrm{E}\ \ln(\chi_f^2) \approx \ln f - \frac{1}{f} - \frac{1}{3f^2} + 2\frac{1}{15f^4} - \cdots \ , \tag{50}$$

and so it holds for the likelihood ratio test statistic from **??**)

$$\mathrm{E}\,(LR_1) \approx 1 + \frac{1}{2(r-1)} \quad . \tag{51}$$

Therefore, we reject the hypothesis $H_0 : \sigma_a^2 = \sigma_b^2$ at the significance level α if

$$LR_1^* = \frac{LR_1}{1 + \dfrac{1}{2(r-1)}} > \chi^2_{1,1-\alpha} \ , \tag{52}$$

and $\chi^2_{1,\gamma}$ denotes the γ–quantile of a χ^2–distribution with one degree of freedom.

If the number of levels of the A– and B–factor are different, i. e. $r \neq s$, we get the following solution of (**??**)

$$\bar{\tau}_4 = M_4 \quad , \quad \bar{\tau}_3 = (s\bar{\tau}_2 - r\bar{\tau}_1)/(s-r) \tag{53}$$

and $\bar{\tau}_1$ and $\bar{\tau}_2$ are solutions of

$$\begin{aligned} (M_1 - \tau_1)(s\tau_2 - r\tau_1)^2 + \tau_1^2(s\tau_2 - r\tau_1 - (s-r)M_3)r(s-1) = 0 \\ (M_2 - \tau_2)(s\tau_2 - r\tau_1)^2 + \tau_2^2(r\tau_1 - s\tau_2 - (r-s)M_3)s(r-1) = 0 \end{aligned} \tag{54}$$

Instead of using a numerical algorithm for computing a solution of (**??**) we use the following approximation. It holds

$$D(\sigma)\Lambda^{-1}K = \begin{bmatrix} 2\sigma_a^4 st/(r-1) \\ -\,2\sigma_b^4 rt/(s-1) \\ 0 \\ 0 \end{bmatrix} + O(r,s) \ , \tag{55}$$

where $O(r,s) \in I\!R^4$ and $\lim O(r,s) = 0$ for $r \to \infty$ and $s \to \infty$.

Thus, instead of solving (**??**) we consider the following system of equations, where we omit the term $O(r,s)$,

$$\begin{array}{rcl}(M_1-\tau_1)+2\lambda\sigma_a^4 st/(r-1) & = & 0\\ (M_2-\tau_2)-2\lambda\sigma_b^4 rt/(s-1) & = & 0\\ (M_3-\tau_3) & = & 0\\ (M_4-\tau_4) & = & 0\\ \frac{1}{s}\tau_1-\frac{1}{r}\tau_2+\left(\frac{1}{r}-\frac{1}{s}\right)\tau_3 & = & 0\end{array} \tag{56}$$

The estimators for τ_3 and τ_4 are now given by

$$\bar{\tau}_3 = M_3 \quad \text{and} \quad \bar{\tau}_4 = M_4\ . \tag{57}$$

With

$$2\lambda = \frac{s-1}{rt}\frac{M_2+\tau_2}{\sigma_b^4} \tag{58}$$

the first equation in (**??**) can be written as

$$(M_1-\tau_1)+\frac{st(s-1)}{rt(r-1)}\frac{\sigma_a^4}{\sigma_b^4}(M_2-\tau_2)=0\ , \tag{59}$$

which under $H_0:\sigma_a^2=\sigma_b^2$ reduces to

$$(M_1-\tau_1)+\frac{s(s-1)}{r(r-1)}(M_2-\tau_2)=0\ . \tag{60}$$

Finally, we yield for τ_1 and τ_2 the estimators

$$\begin{array}{rcl}\bar{\tau}_1 & = & \frac{1}{(r-1)+(s-1)}\left((r-1)M_1+(s-1)\left(\frac{s}{r}M_2+(1-\frac{s}{r})M_3\right)\right)\\ \bar{\tau}_2 & = & \frac{1}{(r-1)+(s-1)}\left((s-1)M_2+(r-1)\left(\frac{r}{s}M_1+(1-\frac{r}{s})M_3\right)\right)\end{array} \tag{61}$$

We note that the approximate solution of (**??**) coincides with the exact solution of (**??**), if the numbers of levels of the A– and B–factor are identical.

Now, we consider the expected value of the 'approximate' likelihood ratio test statistic using the estimators from (**??**) and (**??**) and observe that the last term in both equations on the right hand side of (**??**) is $O(r,s)$ so that estimators from (**??**) can be written as

$$\bar{\tau}_1=\bar{\tau}_2\approx\frac{1}{((r-1)+(s-1))}\left((r-1)M_1+(s-1)M_2\right). \tag{62}$$

Using (**??**) the likelihood ratio test statistic LR can be approximated as

$$LR^{\dagger} = (r-1)\ln\left(\frac{(r-1)M_1+(s-1)M_2}{((r-1)+(s-1))M_1}\right) + (s-1)\ln\left(\frac{(r-1)M_1+(s-1)M_2}{((r-1)+(s-1))M_2}\right) , \tag{63}$$

which under H_0 is appproximately

$$LR^{\dagger} \approx ((r-1)+(s-1))\ln\chi^2(r-1) - (s-1) - (r-1)\ln\chi^2_{r-1} - (s-1)\ln\chi^2_{s-1}$$

$$-((r-1)+(s-1))\ln((r-1)+(s-1)) + (r-1)\ln(r-1) + (s-1)\ln(s-1) \,. \tag{64}$$

With the approximation formula (**??**) we yield for the expected value of the likelihood ratio test statistic

$$\mathrm{E}\,(LR\dagger) \approx 1 + \frac{1}{3(r-1)} + \frac{1}{3(s-1)} - \frac{1}{3((r-1)+(s-1))} \,. \tag{65}$$

Using the inequality

$$\frac{1}{(r-1)+(s-1)} \leq \frac{1}{4}\left(\frac{1}{r-1}+\frac{1}{s-1}\right) \tag{66}$$

the expected value of (**??**) can also be approximated by

$$\mathrm{E}\,(LR^{\dagger}) \approx 1 + \frac{1}{4(r-1)} + \frac{1}{4(s-1)} \,. \tag{67}$$

Therefore, we reject the hypothesis $H_0 : \sigma_a^2 = \sigma_b^2$, if

$$LR^{\ddagger} = LR^{\dagger}/c > \chi^2_{1;1-\alpha} \quad \text{with } c \text{ from } (\mathbf{??}). \tag{68}$$

We note that (**??**) and (**??**) are identical if $r = s$.

In a simulation study, which has been carried out in similar way like the ones in example 4.1, we study the sizes of the proposed tests. We consider the likelihood ratio test using the Newton-Raphson ridge optimization algorithm to maximize the likelihood function under H_0, the 'approximate' likelihood ratio test from (**??**), and the Wald test from (**??**). The results are based on 10 000 simulations and the variance components σ^2_{ab} and σ^2_e have been set equal to one. The results of this simulation study are presented in table 3, where only the results with increasing r and three pairs of (s, t) are reported, because the results with increasing s and different pairs (r, t) are quite similar.

The estimated sizes of the likelihood ratio statistic (LR) exceed for all sample sizes the nominal significance level, but in an acceptable manner. Moreover, the estimated sizes do not depend on the sample sizes on the whole, they are rather homogeneous. Regarding the 'approximate' likelihood statistic ($\mathrm{LR}^{\ddagger}$) we see that the consideration of a factor, who corrects for the expected value of the likelihood ratio test statistic in small smaple sizes, has an important impact on the estimated sizes. In all considered cases, the variation of the estimated significance levels about the nominal significance is rather small. The Wald test mainly produces very conservative results. In our simulation study for $\alpha = 0.01$ the Wald test never rejects the hypothesis, except for $r = 20$. For $\alpha = 0.05$ the estimated sizes are also very small, but sometimes we observe estimated sizes, which seriously exceed 0.05, e. g. the case $r = 3$, $s = 10$, $t = 15$. Thus, the test statistics LR and $\mathrm{LR}^{\ddagger}$, respectively, are more appropriate for testing the hypothesis $H_0 : \sigma_a^2 = \sigma_b^2$ than the Wald test.

References

Abramowitz, M. and I. A. Stegun (1964): *Handbook of Mathematical Functions.* Dover Publications, New York.

Anderson, T. W. (1973): Asymptotically efficient estimation of covariance matrices with linear structure. The Annals of Statistics**1**, 135–141.

Bartlett, M. S. and D. G. Kendall (1946): The statistical analysis of variance–heterogeneity and the logarithmic transformation. *Journal of the Royal Statistical Society B* **8**, 128–138.

Brown, K. G. (1976): Asymptotic behavior of MINQUE–type estimators of variance components. *The Annals of Statistics* **4**, 746–754.

El–Bassiouni, M. Y. and J. Seely (1980): Optimal tests for certain functions of the parameters in a covariance matrix with linear structure. *Sankhya* A **42**, 64–77.

Elpelt, B. (1989): On linear statistical models of commutative quadratic type. *Communication in Statistics – Theory and Methods* **18**, 3407–3450.

Graybill, F. A. (1976): *Theory and Application of the Linear Model.* Duxbury Press, North Scituate, Massachusetts.

Hartung, J. (1981): Nonnegative minimum biased invariant estimation in variance component models. *The Annals of Statistics* **9**, 278–292.

Hartung, J. and B. Voet (1986): Best invariant unbiased estimators for the mean squared error of variance component estimators. *Journal of the American Statistical Association* **81**, 689–691.

Rao, C. R. (1973): *Linear Statistical Inference and Its Applications.* Second Edition, Wiley & Sons, New York.

Seely, J. (1971): Quadratic subspaces and completeness. *Annals of Mathematical Statistics* **42**, 710–721.

Table 1: Estimated size of the likelihood ratio (LR) and Wald test for different sample sizes r, s, t ($\alpha = 0.01$ and $\alpha = 0.05$) in model (??) for testing $H_0 : \sigma_a^2 = \sigma_b^2$.

			$\alpha = 0.01$		$\alpha = 0.05$	
r	s	t	LR	Wald	LR	Wald
3	3	3	0.0133	0.0023	0.0575	0.0136
3	5	6	0.0151	0.0007	0.0654	0.0698
3	8	10	0.0140	0.0734	0.0674	0.1860
4	3	3	0.0124	0.0016	0.0558	0.0155
4	5	6	0.0112	0.0029	0.0579	0.0723
4	8	10	0.0116	0.0732	0.0602	0.1508
5	3	3	0.0110	0.0033	0.0560	0.0163
5	5	6	0.0095	0.0097	0.0521	0.0705
5	8	10	0.0136	0.0653	0.0587	0.1364
6	3	3	0.0125	0.0021	0.0571	0.0243
6	5	6	0.0101	0.0115	0.0540	0.0696
6	8	10	0.0139	0.0546	0.0555	0.1149
7	3	3	0.0091	0.0015	0.0533	0.0224
7	5	6	0.0125	0.0156	0.0542	0.0680
7	8	10	0.0123	0.0526	0.0551	0.1130
8	3	3	0.0092	0.0013	0.0521	0.0247
8	5	6	0.0117	0.0162	0.0509	0.0653
8	8	10	0.0134	0.0522	0.0569	0.1082
9	3	3	0.0112	0.0029	0.0527	0.0258
9	5	6	0.0096	0.0172	0.0569	0.0650
9	8	10	0.0124	0.0481	0.0567	0.0994
10	3	3	0.0102	0.0022	0.0579	0.0315
10	5	6	0.0126	0.0174	0.0509	0.0603
10	8	10	0.0112	0.0451	0.0537	0.0932
20	3	3	0.0105	0.0049	0.0520	0.0402
20	5	6	0.0094	0.0165	0.0519	0.0571
20	8	10	0.0123	0.0321	0.0528	0.0731

Table 2: Estimated power of the likelihood ratio (LR) and Wald test for different values $\delta = \sigma_b^2 - \sigma_a^2$, different sample sizes r, and, $s = 5$, $t = 6$ ($\alpha = 0.05$) in model (**??**) for testing $H_0 : \sigma - a^2 = \sigma_b^2$

		$\delta = \sigma_b^2 - \sigma_a^2$						
r	Test	0	0.1	0.25	0.5	0.75	1	2
3	LR	0.0654	0.0699	0.0825	0.0986	0.1119	0.1224	0.1516
	Wald	0.0698	0.0877	0.1270	0.2030	0.2638	0.3090	0.4114
4	LR	0.0579	0.0625	0.0816	0.1119	0.1329	0.1493	0.2045
	Wald	0.0723	0.0962	0.1461	0.2345	0.2957	0.3438	0.4656
5	LR	0.0521	0.0609	0.0843	0.1292	0.1604	0.1848	0.2553
	Wald	0.0705	0.1019	0.1633	0.2589	0.3242	0.3779	0.5180
6	LR	0.0540	0.0662	0.0970	0.1505	0.1856	0.2210	0.3149
	Wald	0.0696	0.1062	0.1751	0.2823	0.3593	0.4202	0.5712
7	LR	0.0542	0.0738	0.1130	0.1652	0.2132	0.2524	0.3681
	Wald	0.0680	0.1099	0.1891	0.2948	0.3811	0.4496	0.6128
8	LR	0.0509	0.0701	0.1142	0.1814	0.2362	0.2850	0.4260
	Wald	0.0653	0.1071	0.1997	0.3203	0.4146	0.4870	0.6541
9	LR	0.0569	0.0756	0.1284	0.2079	0.2718	0.3267	0.4774
	Wald	0.0650	0.1119	0.2043	0.3430	0.4433	0.5138	0.6913
10	LR	0.0509	0.0733	0.1286	0.2185	0.2926	0.3528	0.5206
	Wald	0.0603	0.1115	0.2023	0.3559	0.4586	0.5320	0.7222
20	LR	0.0519	0.1019	0.2218	0.3798	0.5020	0.5951	0.8175
	Wald	0.0571	0.1417	0.3039	0.4956	0.6237	0.7224	0.8963

Table 3: Estimated size of the likelihood ratio (LR), the approximate likelihood ratio ($LR^{\ddagger}$), and Wald test for different sample sizes r, s, t ($\alpha = 0.01$ and $\alpha = 0.05$) in model (**??**) for testing $H_0 : \sigma_a^2 = \sigma_b^2$.

			$\alpha = 0.01$			$\alpha = 0.05$		
r	s	t	LR	$LR^{\ddagger}$	Wald	LR	$LR^{\ddagger}$	Wald
3	5	6	0.0144	0.0082	0	0.0716	0.0466	0
3	8	10	0.0158	0.0094	0	0.0720	0.0501	0.0616
3	10	15	0.0167	0.0083	0	0.0689	0.0487	0.1229
4	5	6	0.0163	0.0100	0	0.0682	0.0505	0
4	8	10	0.0165	0.0098	0	0.0654	0.0508	0.0265
4	10	15	0.0146	0.0102	0	0.0658	0.0504	0.069
5	5	6	0.0141	0.0089	0	0.0636	0.0502	0
5	8	10	0.0141	0.0098	0	0.0626	0.0475	0.0114
5	10	15	0.0133	0.0100	0	0.0604	0.0489	0.0383
6	5	6	0.0131	0.0095	0	0.0596	0.0461	0
6	8	10	0.0141	0.0104	0	0.0618	0.0506	0.0057
6	10	15	0.0122	0.0095	0	0.0580	0.0479	0.0228
7	5	6	0.0137	0.0090	0	0.0625	0.0514	0.0011
7	8	10	0.0147	0.0110	0	0.0584	0.0493	0.0026
7	10	15	0.0134	0.0109	0	0.0588	0.0502	0.0152
8	5	6	0.0144	0.0101	0	0.0648	0.0505	0.0113
8	8	10	0.0120	0.0095	0	0.0586	0.0503	0.0018
8	10	15	0.0145	0.0116	0	0.0604	0.0543	0.0126
9	5	6	0.0153	0.0106	0	0.0598	0.0495	0.0272
9	8	10	0.0131	0.0100	0	0.0582	0.0489	0.0061
9	10	15	0.0118	0.0091	0	0.0559	0.0487	0.0092
10	5	6	0.0151	0.0112	0	0.0620	0.0517	0.0413
10	8	10	0.0125	0.0098	0	0.0586	0.0522	0.0115
10	10	15	0.0111	0.0095	0	0.0555	0.0470	0.0097
20	5	6	0.0154	0.0115	0.046	0.0599	0.0524	0.1312
20	8	10	0.0123	0.0105	0.012	0.0585	0.0515	0.0637
20	10	15	0.0133	0.0082	0.005	0.0642	0.0483	0.0466